ESE 650: Learning in Robotics
Lecture 7

Lecturer:

Nikolay Atanasov: atanasov@seas.upenn.edu

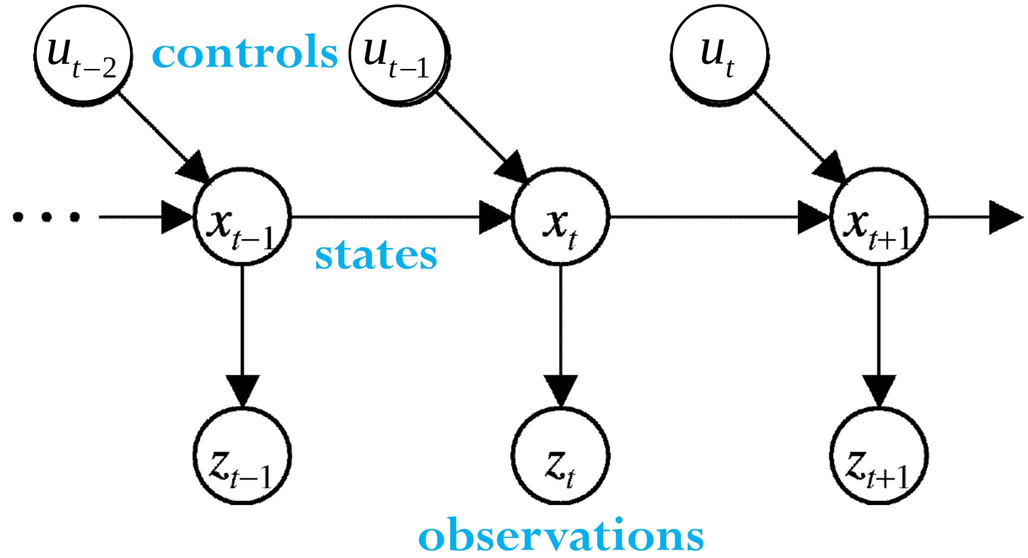
Teaching Assistants:

Jinwook Huh: jinwookh@seas.upenn.edu

Heejin Jeong: heejinj@seas.upenn.edu

Kelsey Saulnier: saulnier@seas.upenn.edu

Bayes Filter (summary)



Performs probabilistic tracking of

$$p_{t|t}(x_t) := p(x_t \mid z_{0:t}, u_{0:t-1})$$

$$p_{t+1|t}(x_{t+1}) := p(x_{t+1} \mid z_{0:t}, u_{0:t})$$

Posterior Distribution:

$$p_{t+1|t+1}(x_{t+1}) = \eta\, p_h(z_{t+1} \mid x_{t+1}) \int p_a(x_{t+1} \mid x_t, u_t) p_{t|t}(x_t) dx_t$$

Normalization constant: $\eta = \dfrac{1}{p(z_{t+1} \mid z_{0:t}, u_{0:t})}$

Joint Distribution:

prior

observation model

motion model

$$p\left(x_{0:T}, z_{0:T}, u_{0:T-1}\right) = p_{0|0}(x_0) \prod_{t=0}^{T} p_h(z_k \mid x_k) \prod_{t=0}^{T} p_a(x_k \mid x_{k-1}, u_{k-1})$$

Kalman Filter

- Estimates the state x_t of a discrete-time controlled process that is governed by a linear stochastic difference equation (motion model) using measurements z_t obtained from a linear stochastic transformation of the state

- A **Kalman filter** is a Bayes filter for which:
  - The prior pdf is Gaussian
  - The motion model is linear in the state and affected by Gaussian noise
  - The observation model is linear in the state and affected by Gaussian noise
  - The process and measurement noises are independent of each other and the state

- Motion model: $x_{t+1} = a(x_t, u_t, w_t) := Ax_t + Bu_t + w_t \qquad w_t \sim \mathcal{N}(0, W)$

  $$x_{t+1} \mid x_t, u_t \sim \mathcal{N}(Ax_t + Bu_t, W)$$

- Observation model: $z_t = h(x_t, v_t) := Hx_t + v_t \qquad v_t \sim \mathcal{N}(0, V)$

  $$z_t \mid x_t \sim \mathcal{N}(Hx_t, V)$$

Kalman Filter Prediction

$$p_{t+1|t}(x) = \int p_a(x \mid s, u_t)\, p_{t|t}(s) ds = \int \phi(x; As + Bu_t, W)\phi(s; \mu_{t|t}, \Sigma_{t|t}) ds$$

$$= \eta \int \exp\left\{ -\frac{1}{2}(x - As - Bu_t)^T W^{-1}(x - As - Bu_t) \right\}$$

$$\exp\left\{ -\frac{1}{2}(s - \mu_{t|t})^T \Sigma_{t|t}^{-1}(s - \mu_{t|t}) \right\} ds$$

$$= \eta \int \exp\left\{ -\frac{1}{2}\left(s^T \left(A^T W^{-1} A + \Sigma_{t|t}^{-1} \right) s - 2\left(\Sigma_{t|t}^{-1}\mu_{t|t} + A^T W^{-1}(x - Bu_t) \right)^T s + \ldots \right) \right\} ds$$

$$= \phi\left(x; A\mu_{t|t} + Bu_t, A\Sigma_{t|t}A^T + W \right)$$

Square Completion & Matrix Inversion Lemma

$$\boxed{p_{t+1|t}(x) = \int \phi(x; As + Bu_t, W)\phi(s; \mu_{t|t}, \Sigma_{t|t}) ds = \phi(x; A\mu_{t|t} + Bu_t, A\Sigma_{t|t}A^T + W)}$$

Kalman Filter Prediction (Alternative)

- Motion model with uncertain state:

$$x_{t+1} = Ax_t + Bu_t + w_t \qquad w_t \sim \mathcal{N}(0, W), \qquad x_t \sim \mathcal{N}(\mu_{t|t}, \Sigma_{t|t})$$

- Suppose we *know* that the distribution of the predicted state is Gaussian: $x_{t+1} \sim \mathcal{N}(\mu_{t+1|t}, \Sigma_{t+1|t})$, then we just need to compute its mean and covariance:

$$\mu_{t+1|t} = \mathbb{E}\left[Ax_t + Bu_t + w_t\right] = A\mathbb{E}\left[x_t\right] + Bu_t + \mathbb{E}\left[w_t\right] = A\mu_{t|t} + Bu_t$$

$$\begin{aligned}
\mathbb{E}\left[x_{t+1}x_{t+1}^T\right] &= \mathbb{E}\left[(Ax_t + Bu_t + w_t)(Ax_t + Bu_t + w_t)^T\right] \\
&= A\mathbb{E}\left[x_t x_t^T\right]A^T + A\mathbb{E}\left[x_t\right]u_t^T B^T + A\mathbb{E}\left[x_t w_t^T\right] \\
&\quad + Bu_t\mathbb{E}\left[x_t^T\right]A + Bu_t u_t^T B^T + Bu_t\mathbb{E}\left[w_t^T\right] \\
&\quad + \mathbb{E}\left[w_t x_t^T\right]A^T + \mathbb{E}\left[w_t\right]u_t^T B^T + \mathbb{E}\left[w_t w_t^T\right] \\
&= A\left(\Sigma_{t|t} + \mu_{t|t}\mu_{t|t}^T\right)A^T + A\mu_{t|t}u_t^T B + Bu_t\mu_{t|t}^T A + Bu_t u_t^T B^T + W
\end{aligned}$$

Kalman Filter Update

$$p_{t+1|t+1}(x) = \frac{p_h(z_{t+1} \mid x)\, p_{t+1|t}(x)}{p(z_{t+1} \mid z_{0:t}, u_{0:t})} = \frac{\phi(z_{t+1}; Hx, V)\phi(x; \mu_{t+1|t}, \Sigma_{t+1|t})}{\int \phi(z_{t+1}; Hs, V)\phi(s; \mu_{t+1|t}, \Sigma_{t+1|t})ds}$$

$$= \eta \exp\left\{-\frac{1}{2}(z_{t+1} - Hx)^T V^{-1}(z_{t+1} - Hx)\right\} \exp\left\{-\frac{1}{2}(x - \mu_{t+1|t})^T \Sigma_{t+1|t}^{-1}(x - \mu_{t+1|t})\right\}$$

$$= \eta \exp\left\{-\frac{1}{2}\left(x^T \left(H^T V^{-1} H + \Sigma_{t+1|t}^{-1}\right) x + \left(H^T V^{-1} z_{t+1} + \Sigma_{t+1|t}^{-1}\mu_{t+1|t}\right)^T x + \ldots\right)\right\}$$

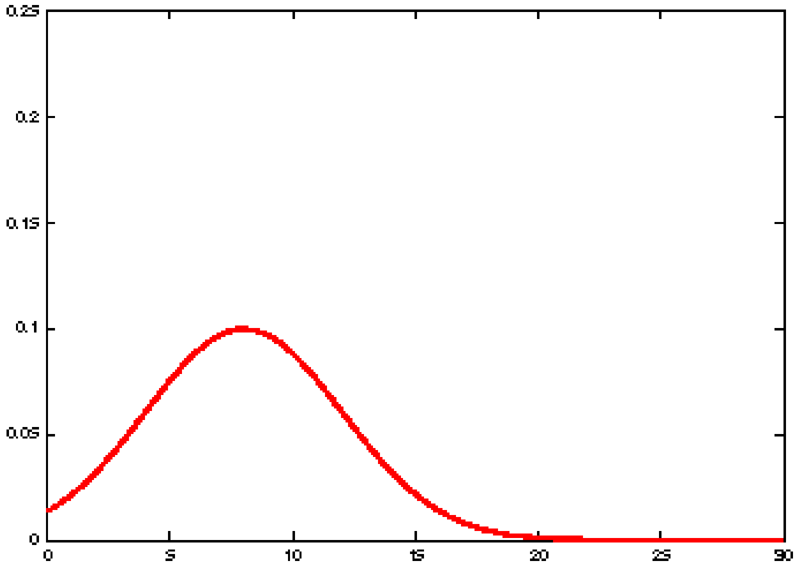
$$= \phi\left(x; \left(H^T V^{-1} H + \Sigma_{t+1|t}^{-1}\right)^{-1}\left(H^T V^{-1} z_{t+1} + \Sigma_{t+1|t}^{-1}\mu_{t+1|t}\right), \left(H^T V^{-1} H + \Sigma_{t+1|t}^{-1}\right)^{-1}\right)$$

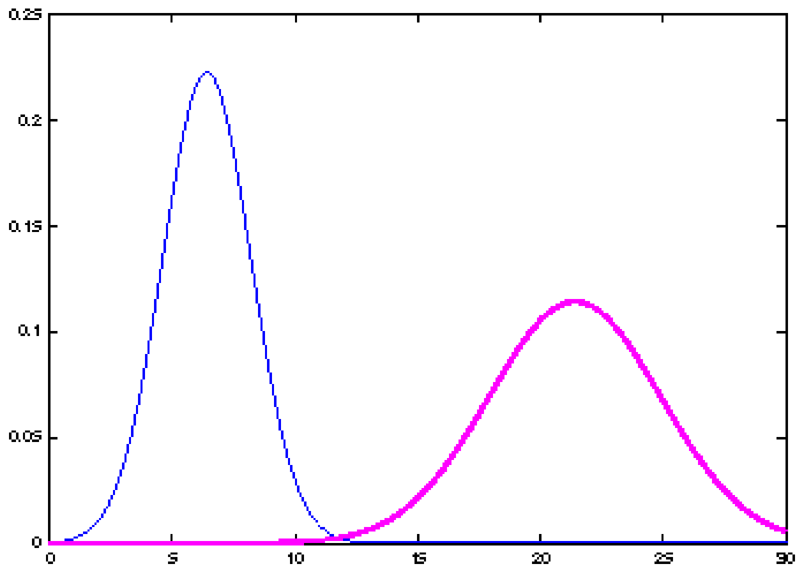
$$\boxed{= \phi\left(x; \mu_{t+1|t} + K_{t+1|t}(z_{t+1} - H\mu_{t+1|t}), (I - K_{t+1}H)\Sigma_{t+1|t}\right)}$$

Square Completion & Matrix Inversion Lemma

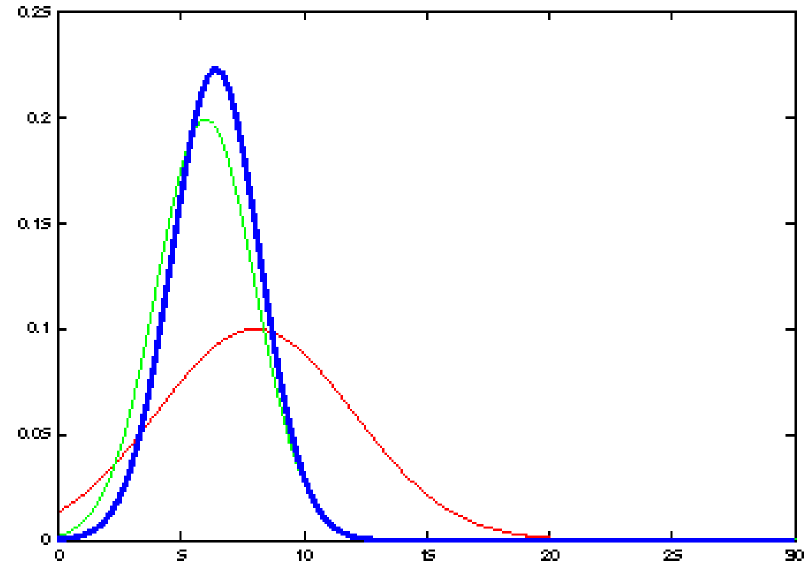
- **Kalman gain:** $K_{t+1|t} := \Sigma_{t+1|t} H^T \left(H \Sigma_{t+1|t} H^T + V\right)^{-1}$
- **Square completion:** $\frac{1}{2}x^T A x + b^T x + c = \frac{1}{2}\left(x + A^{-1}b\right)^T A\left(x + A^{-1}b\right) + c - \frac{1}{2}b^T A^{-1} b$

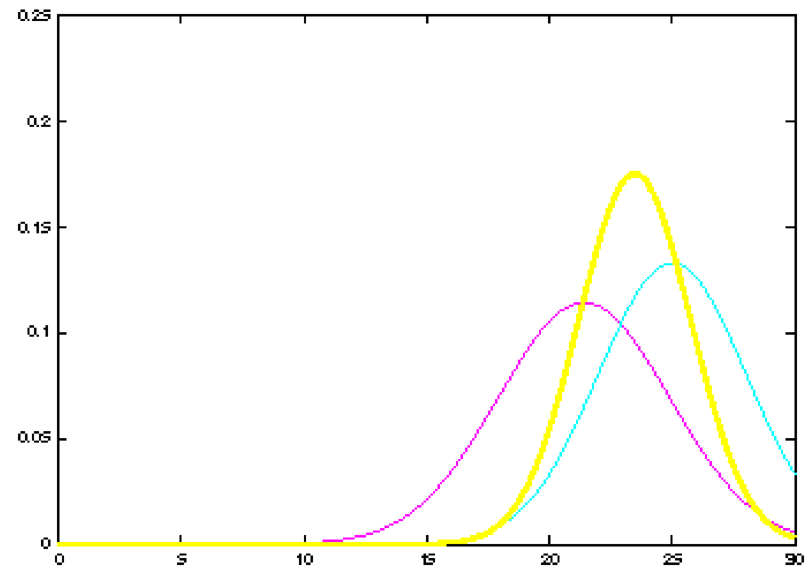
Prediction step





Update step





Kalman Filter (discrete time)

Motion model: $$x_{t+1} = Ax_t + Bu_t + \mathcal{N}(0, W)$$

Observation model: $$z_t = Hx_t + \mathcal{N}(0, V)$$

Prior: $$x_t \mid z_{0:t}, u_{0:t-1} \sim \mathcal{N}(\mu_{t|t}, \Sigma_{t|t})$$

Prediction:
$$\mu_{t+1|t} = A\mu_{t|t} + Bu_t$$
$$\Sigma_{t+1|t} = A\Sigma_{t|t}A^T + W$$

Update:
$$\mu_{t+1|t+1} = \mu_{t+1|t} + K_{t+1|t}(z_{t+1} - H\mu_{t+1|t})$$
$$\Sigma_{t+1|t+1} = (I - K_{t+1}H)\Sigma_{t+1|t}$$

Kalman Gain: $$K_{t+1|t} := \Sigma_{t+1|t}H^T\left(H\Sigma_{t+1|t}H^T + V\right)^{-1}$$

Kalman-Bucy Filter (continuous time)

Motion model: $\dot{x}(t) = Ax(t) + Bu(t) + w(t)$

Observation model: $z(t) = Hx(t) + v(t)$

Prior: $x(0) \sim \mathcal{N}(\mu(0), \Sigma(0))$

Mean: $\dot{\mu}(t) = A\mu(t) + Bu(t) + K(t)\left(z(t) - H\mu(t)\right)$

Covariance: $\dot{\Sigma}(t) = A\Sigma(t) + \Sigma(t)A^T + W - K(t)VK^T(t)$

Kalman Gain: $K(t) = \Sigma(t)H^T V^{-1}$

EM for Kalman Filtering

- If we are uncertain about the parameters (A, B, W) of the motion model or the parameters (H, V) of the observation model, we can learn them using EM!

- Given data $\{z_{0:T}, u_{0:T-1}\}$, apply the EM algorithm with hidden variable x_T
  - **E step**: Given initial parameter estimates $\theta^{(i)} \coloneqq \{A^{(i)}, B^{(i)}, W^{(i)}, H^{(i)}, V^{(i)}\}$ calculate the likelihood of the hidden variable
  - **M step**: Optimize the parameters to obtain $\theta^{(i+1)}$ which better explains the posterior distribution over x_T

Kalman Filter (summary)

- **Highly efficient**: polynomial in measurement dimensionality p and state dimensionality n: $\boxed{O(p^{2.376} + n^2)}$
- **Optimal** for linear Gaussian systems!
- Most robotics systems are **nonlinear**!

Nonlinear Kalman Filtering

- A **nonlinear Kalman filter** is a Bayes filter for which:
  - The prior pdf is Gaussian
  - The motion model is ~~linear in the state and~~ affected by Gaussian noise
  - The observation model is ~~linear in the state and~~ affected by Gaussian noise
  - The process and measurement noises are independent of each other and the state
  - The posterior pdf is **forced to be Gaussian via approximation**

- Motion model: $x_{t+1} = a(x_t, u_t, w_t) \qquad w_t \sim \mathcal{N}(0, W)$

- Observation model: $z_t = h(x_t, v_t) \qquad v_t \sim \mathcal{N}(0, V)$

- The **main challenge** is that the predicted and updated pdfs can no longer be evaluated in closed form because the posterior state and measurement distributions are not necessarily Gaussian

- However, we can force the predicted and updated pdfs to be Gaussian via **moment matching**, i.e., by evaluating the first and second moments of the actual non-Gaussian pdfs and approximating them with a Gaussian distribution with the same first and second moments

Nonlinear Kalman Filter

- **Prior:** $x_t \mid z_{0:t}, u_{0:t-1} \sim \mathcal{N}\left(0, \mu_{t|t}, \Sigma_{t|t}\right)$
- **Prediction step**: $x_{t+1} \mid z_{0:t}, u_{0:t} \sim \mathcal{N}\left(0, \mu_{t+1|t}, \Sigma_{t+1|t}\right)$

$$p_{t+1|t}(x) = \int p_a(x \mid s, u_t) p_{t|t}(s) ds = \int p_a(x \mid s, u_t) \phi\left(s; \mu_{t|t}, \Sigma_{t|t}\right) ds$$

$$\mu_{t+t|t} := \mathbb{E}\left[x_{t+1} \mid z_{0:t}, u_{0:t}\right] = \int x p_{t+1|t}(x) dx = \int x \left[\int p_a(x \mid s, u_t) \phi\left(s; \mu_{t|t}, \Sigma_{t|t}\right) ds\right] dx$$

$$\Sigma_{t+1|t} := \mathbb{E}\left[x_{t+1} x_{t+1}^T \mid z_{0:t}, u_{0:t}\right] - \mu_{t+t|t} \mu_{t+1|t}^T = \int x x^T p_{t+1|t}(x) dx - \mu_{t+t|t} \mu_{t+1|t}^T$$

- **Update step**: $x_{t+1} \mid z_{0:t+1}, u_{0:t} \sim \mathcal{N}\left(0, \mu_{t+1|t+1}, \Sigma_{t+1|t+1}\right)$

$$p_{t+1|t+1}(x) = \frac{p_h(z_{t+1} \mid x) p_{t+1|t}(x)}{\int p_h(z_{t+1} \mid s) p_{t+1|t}(s) ds}$$

$$\mu_{t+1|t+1} := \mathbb{E}\left[x_{t+1} \mid z_{0:t+1}, u_{0:t}\right] = \int x p_{t+1|t+1}(x) dx$$

$$\Sigma_{t+1|t+1} := \mathbb{E}\left[x_{t+1} x_{t+1}^T \mid z_{0:t+1}, u_{0:t}\right] - \mu_{t+t|t+1} \mu_{t+1|t+1}^T = \int x x^T p_{t+1|t+1}(x) dx - \mu_{t+t|t+!} \mu_{t+1|t+1}^T$$

Moment Matching

- The **moment matching** Gaussian approximation to the joint distribution of x and a transformed random variable $z = h(x, v)$ where $x \sim \mathcal{N}(\mu, \Sigma)$ and $v \sim \mathcal{N}(0, V)$ is:

$$\begin{pmatrix} x \\ z \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \mu \\ m \end{pmatrix}, \begin{bmatrix} \Sigma & C \\ C^T & S \end{bmatrix} \right)$$

$$m := \int\int h(x,v)\phi(x;\mu,\Sigma)\,\phi(v;0,V)dxdv$$

$$S := \int\int \big(h(x,v)-m\big)\big(h(x,v)-m\big)^T \phi(x;\mu,\Sigma)\,\phi(v;0,V)dxdv$$

$$C := \int\int \big(x-\mu\big)\big(h(x,v)-m\big)^T \phi(x;\mu,\Sigma)\,\phi(v;0,V)dxdv$$

- **Update step**: $$p_{t+1|t+1}(x) = \frac{p_h(z_{t+1} \mid x)\, p_{t+1|t}(x)}{\int p_h(z_{t+1} \mid s)\, p_{t|t}(s)ds} \approx \phi\left(x; \mu_{t+1|t+1}, \Sigma_{t+1|t+1}\right)$$

$$\boxed{\begin{aligned} \mu_{t+1|t+1} &:= \mu_{t+1|t} + K_{t+1|t}\left(z_{t+1} - m_{t+1|t}\right) \\ \Sigma_{t+1|t+1} &:= \Sigma_{t+1|t} - K_{t+1|t} S_{t+1|t} K_{t+1|t} \\ K_{t+1|t} &:= C_{t+1|t} S_{t+1|t}^{-1} \end{aligned}}$$

Nonlinear KF for Additive Noise (prediction)

- **Motion model:** $x_{t+1} = a(x_t, u_t) + w_t \qquad x_{t+1} \mid x_t, u_t \sim \mathcal{N}\left(a(x_t, u_t), W\right)$

- **Prediction step**:

$$p_{t+1|t}(x) = \int p_a(x \mid s, u_t)\, p_{t|t}(s)ds = \int \phi\left(x; a(s,u_t), W\right)\phi\left(s; \mu_{t|t}, \Sigma_{t|t}\right)ds = \phi\left(x; \mu_{t+1|t}, \Sigma_{t+1|t}\right)$$

$$\begin{aligned}\mu_{t+t|t} &:= \int x p_{t+1|t}(x)dx = \int\left[\int x\phi\left(x; a(s,u_t), W\right)dx\right]\phi\left(s; \mu_{t|t}, \Sigma_{t|t}\right)ds \\ &= \underline{\int a(s,u_t)\phi\left(s; \mu_{t|t}, \Sigma_{t|t}\right)ds}\end{aligned}$$

$$\begin{aligned}\Sigma_{t+1|t} &:= \int xx^T p_{t+1|t}(x)dx - \mu_{t+t|t}\mu_{t+1|t}^T = \int\left[\int xx^T\phi\left(x; a(s,u_t), W\right)dx\right]\phi\left(s; \mu_{t|t}, \Sigma_{t|t}\right)ds - \mu_{t+t|t}\mu_{t+1|t}^T \\ &= \underline{\int a(s,u_t)\left(a(s,u_t)\right)^T \phi\left(s; \mu_{t|t}, \Sigma_{t|t}\right)ds + W - \mu_{t+t|t}\mu_{t+1|t}^T}\end{aligned}$$

Nonlinear KF for Additive Noise (summary)

- **Motion model:** $x_{t+1} = a(x_t, u_t) + w_t \qquad x_{t+1} \mid x_t, u_t \sim \mathcal{N}\left(a(x_t, u_t), W\right)$
- **Measurement model:** $z_t = h(x_t) + v_t \qquad z_t \mid x_t \sim \mathcal{N}\left(h(x_t), V\right)$
- **Prediction step:**

$$\mu_{t+t|t} = \int a(s, u_t)\phi\left(s; \mu_{t|t}, \Sigma_{t|t}\right) ds$$

$$\Sigma_{t+1|t} = \int a(s, u_t)\left(a(s, u_t)\right)^T \phi\left(s; \mu_{t|t}, \Sigma_{t|t}\right) ds + W - \mu_{t+t|t}\mu_{t+1|t}^T$$

- **Update step**:

$$\mu_{t+1|t+1} := \mu_{t+1|t} + K_{t+1|t}\left(z_{t+1} - m_{t+1|t}\right)$$

$$\Sigma_{t+1|t+1} := \Sigma_{t+1|t} - K_{t+1|t} S_{t+1|t} K_{t+1|t}$$

$$K_{t+1|t} := C_{t+1|t} S_{t+1|t}^{-1}$$

$$m_{t+1|t} := \int h(x)\phi\left(x; \mu_{t+1|t}, \Sigma_{t+1|t}\right) dx$$

$$S_{t+1|t} := \int \left(h(x) - m_{t+1|t}\right)\left(h(x) - m_{t+1|t}\right)^T \phi\left(x; \mu_{t+1|t}, \Sigma_{t+1|t}\right) dx + V$$

$$C_{t+1|t} := \int \left(x - \mu_{t+1|t}\right)\left(h(x) - m_{t+1|t}\right)^T \phi\left(x; \mu_{t+1|t}, \Sigma_{t+1|t}\right) dx$$

Extended Kalman Filter (prediction)

- **Idea**: use first order Taylor series to approximate the nonlinear motion and measurement models via linear functions! Then, the integrals needed for the Nonlinear Kalman Filter can then be evaluated!

- **Prior pdf**: $x_t \mid z_{0:t}, u_{0:t-1} \sim \mathcal{N}\left(0, \mu_{t|t}, \Sigma_{t|t}\right)$

- **Motion model**: $x_{t+1} = a(x_t, u_t) + w_t \qquad w_t \sim \mathcal{N}(0, W)$

$$a(x_t, u_t) \approx a(\mu_{t|t}, u_t) + \left[\frac{\partial a}{\partial x}\left(\mu_{t|t}, u_t\right)\right](x_t - \mu_{t|t}) \qquad A_t := \frac{\partial a}{\partial x}\left(\mu_{t|t}, u_t\right)$$

- **Prediction step:**

$$\mu_{t+t|t} = \int a(s, u_t)\phi\left(s; \mu_{t|t}, \Sigma_{t|t}\right) ds \approx a(\mu_{t|t}, u_t) + A_t\left(\int s\phi\left(s; \mu_{t|t}, \Sigma_{t|t}\right) ds - \mu_{t|t}\right) = a(\mu_{t|t}, u_t)$$

$$\begin{aligned}\Sigma_{t+1|t} &= \int a(s, u_t)\left(a(s, u_t)\right)^T \phi\left(s; \mu_{t|t}, \Sigma_{t|t}\right) ds + W - \mu_{t+t|t}\mu_{t+1|t}^T \\ &\approx a(\mu_{t|t}, u_t)a^T(\mu_{t|t}, u_t) + A_t \int (s - \mu_{t|t})\phi\left(s; \mu_{t|t}, \Sigma_{t|t}\right) ds a^T(\mu_{t|t}, u_t) \\ &\quad + a(\mu_{t|t}, u_t)\int (s - \mu_{t|t})^T \phi\left(s; \mu_{t|t}, \Sigma_{t|t}\right) ds A_t^T \\ &\quad + A_t \int (s - \mu_{t|t})(s - \mu_{t|t})^T \phi\left(s; \mu_{t|t}, \Sigma_{t|t}\right) ds A_t^T + W - \mu_{t+t|t}\mu_{t+1|t}^T = A_t \Sigma_{t|t} A_t^T + W\end{aligned}$$

Extended Kalman Filter (update)

- **Predicted pdf**: $x_{t+1} \mid z_{0:t}, u_{0:t} \sim \mathcal{N}\left(0, \mu_{t+1|t}, \Sigma_{t+1|t}\right)$
- **Measurement model**: $z_{t+1} = h(x_{t+1}) + v_{t+1} \qquad v_{t+1} \sim \mathcal{N}(0, V)$

$$h(x_{t+1}) \approx h(\mu_{t+1|t}) + \left[\frac{\partial h}{\partial x}\left(\mu_{t+1|t}\right)\right](x_{t+1} - \mu_{t+1|t}) \qquad H_{t+1} := \frac{\partial h}{\partial x}\left(\mu_{t+1|t}\right)$$

- **Update step**:

$$\mu_{t+1|t+1} := \mu_{t+1|t} + K_{t+1|t}\left(z_{t+1} - m_{t+1|t}\right)$$

$$\Sigma_{t+1|t+1} := \Sigma_{t+1|t} - K_{t+1|t} S_{t+1|t} K_{t+1|t}$$

$$K_{t+1|t} := C_{t+1|t} S_{t+1|t}^{-1}$$

$$m_{t+1|t} := \int h(x)\phi\left(x; \mu_{t+1|t}, \Sigma_{t+1|t}\right)dx \approx h(\mu_{t+1|t})$$

$$S_{t+1|t} := \int \left(h(x) - m_{t+1|t}\right)\left(h(x) - m_{t+1|t}\right)^T \phi\left(x; \mu_{t+1|t}, \Sigma_{t+1|t}\right)dx + V \approx H_{t+1}\Sigma_{t+1|t}H_{t+1}^T + V$$

$$C_{t+1|t} := \int \left(x - \mu_{t+1|t}\right)\left(h(x) - m_{t+1|t}\right)^T \phi\left(x; \mu_{t+1|t}, \Sigma_{t+1|t}\right)dx \approx \Sigma_{t+1|t}H_{t+1}^T$$

Extended Kalman Filter (summary)

Prior: $x_t \mid z_{0:t}, u_{0:t-1} \sim \mathcal{N}\left(0, \mu_{t|t}, \Sigma_{t|t}\right)$

Motion Model: $x_{t+1} = a(x_t, u_t) + w_t \qquad w_t \sim \mathcal{N}(0, W)$

$$a(x_t, u_t) \approx a(\mu_{t|t}, u_t) + A_t(x_t - \mu_{t|t}) \qquad A_t := \frac{\partial a}{\partial x}\left(\mu_{t|t}, u_t\right)$$

Measurement Model: $z_{t+1} = h(x_{t+1}) + v_{t+1} \qquad v_{t+1} \sim \mathcal{N}(0, V)$

$$h(x_{t+1}) \approx h(\mu_{t+1|t}) + H_{t+1}\left(x_{t+1} - \mu_{t+1|t}\right) \qquad H_{t+1} := \frac{\partial h}{\partial x}\left(\mu_{t+1|t}\right)$$

Prediction:

$$\mu_{t+1|t} = a(\mu_{t|t}, u_t)$$
$$\Sigma_{t+1|t} = A_t \Sigma_{t|t} A_t^T + W$$

Update:

$$\mu_{t+1|t+1} = \mu_{t+1|t} + K_{t+1|t}\left(z_{t+1} - h(\mu_{t+1|t})\right)$$
$$\Sigma_{t+1|t+1} = (I - K_{t+1} H_{t+1})\Sigma_{t+1|t}$$

Kalman Gain: $K_{t+1|t} := \Sigma_{t+1|t} H_{t+1}^T \left(H_{t+1} \Sigma_{t+1|t} H_{t+1}^T + V\right)^{-1}$

Same as KF but using (A_t, W) and (H_t, V)

Unscented Transform

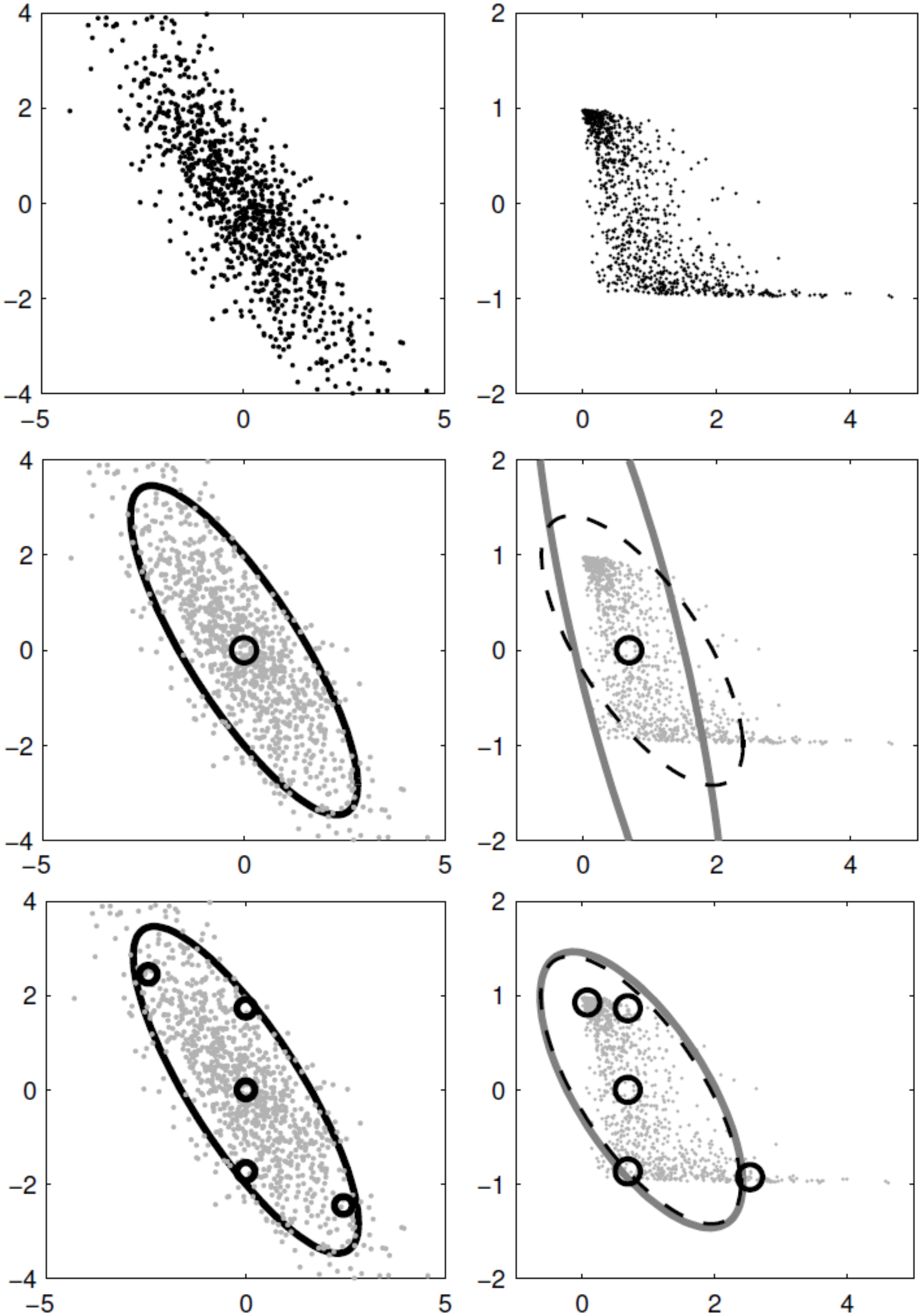
- The **unscented transform** (Julier et al., 1995, 2000) is a numerical method for approximating the joint distribution of random variables x and z defined as

$$z = h(x) \qquad x \sim \mathcal{N}(\mu, \Sigma)$$

- **Idea**: deterministically choose a fixed number of **sigma points** that capture the mean and covariance of the original distribution of x exactly

- Instead of linearizing the function, propagate the sigma points through the nonlinear function and then estimate the mean and covariance of the transformed variable from them

- This resembles Monte Carlo approximation but the sigma points are selected **deterministically**



Applying a nonlinear transformation to the random variable on the left results in a new random variable (right)

EKF: Linearization-based approximation to the transformation formed by calculating the curvature at the mean (solid) and true covariance (dashed).

UKF: Unscented transform approximation to the transformation formed by propagating sigma points through the nonlinearity and estimating the covariance from them (solid) and true covariance (dashed).

- The **unscented transform** Gaussian approximation to the joint distribution of x and a transformed random variable $z = h(x, v)$ where $x \sim \mathcal{N}(\mu, \Sigma)$ and $v \sim \mathcal{N}(0, V)$ is:

$$\begin{pmatrix} x \\ z \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \mu \\ m_U \end{pmatrix}, \begin{bmatrix} \Sigma & C_U \\ C_U^T & S_U \end{bmatrix} \right)$$

- Form a set of $\mathbf{2n+1}$ sigma points using the i-th columns of the square root $\sqrt{\Sigma}$ of the covariance $\Sigma = [\sqrt{\Sigma}][\sqrt{\Sigma}]^T$ (**note**: $\sqrt{\Sigma}$ is lower-triangular and can be obtained, e.g., via **Cholesky factorization**) as follows:

$$\mathcal{X}^{(0)} = \mu$$

$$\mathcal{X}^{(i)} = \mu \pm \sqrt{n+\lambda}\left[\sqrt{\Sigma}\right]_i \qquad i = 1, \ldots, n$$

$$\lambda := \alpha^2 (n+k) - n \qquad (\alpha, k):\ \text{determine sigma points spread, e.g., } \alpha = \sqrt{2}, k = 0$$

- Estimate the mean and covariance of $z = h(x)$ from the transformed sigma points:

$$m_U = \sum_{i=0}^{2n} W_i^{(m)} h\left(\mathcal{X}^{(i)}\right) \qquad W_0^{(m)} = \frac{\lambda}{n+\lambda} \qquad W_i^{(m)} = \frac{1}{2(n+\lambda)} \qquad i = 1, \ldots, 2n$$

$$S_U = \sum_{i=0}^{2n} W_i^{(c)} \left(h\left(\mathcal{X}^{(i)}\right) - m_U\right)\left(h\left(\mathcal{X}^{(i)}\right) - m_U\right)^T + V \qquad W_0^{(c)} = \frac{\lambda}{n+\lambda} + \left(1 - \alpha^2 + \beta\right)$$

$$C_U = \sum_{i=0}^{2n} W_i^{(c)} \left(\mathcal{X}^{(i)} - \mu\right)\left(h\left(\mathcal{X}^{(i)}\right) - m_U\right)^T \qquad W_i^{(c)} = \frac{1}{2(n+\lambda)} \qquad i = 1, \ldots, 2n$$

Unscented Kalman Filter (summary)

Prior: $x_t \mid z_{0:t}, u_{0:t-1} \sim \mathcal{N}\left(0, \mu_{t|t}, \Sigma_{t|t}\right)$

Motion Model: $x_{t+1} = a(x_t, u_t) + w_t \qquad w_t \sim \mathcal{N}(0, W)$

Measurement Model: $z_{t+1} = h(x_{t+1}) + v_{t+1} \qquad v_{t+1} \sim \mathcal{N}(0, V)$

Prediction:

$$\mu_{t+1|t} = \sum_{i=0}^{2n} W_i^{(m)} a\left(\mathcal{X}_{t|t}^{(i)}, u_t\right) \qquad \mathcal{X}_{t|t}^{(0)} = \mu_{t|t} \qquad \mathcal{X}_{t|t}^{(i)} = \mu_{t|t} \pm \sqrt{n+\lambda}\left[\sqrt{\Sigma_{t|t}}\right]_i$$

$$\Sigma_{t+1|t} = \sum_{i=0}^{2n} W_i^{(c)} \left(a\left(\mathcal{X}_{t|t}^{(i)}, u_t\right) - \mu_{t+1|t}\right)\left(a\left(\mathcal{X}_{t|t}^{(i)}, u_t\right) - \mu_{t+1|t}\right)^T + W$$

Update:

$$\mu_{t+1|t+1} = \mu_{t+1|t} + K_{t+1|t}\left(z_{t+1} - m_{t+1|t}\right)$$

$$\Sigma_{t+1|t+1} = \Sigma_{t+1|t} - K_{t+1|t} S_{t+1|t} K_{t+1|t}^T$$

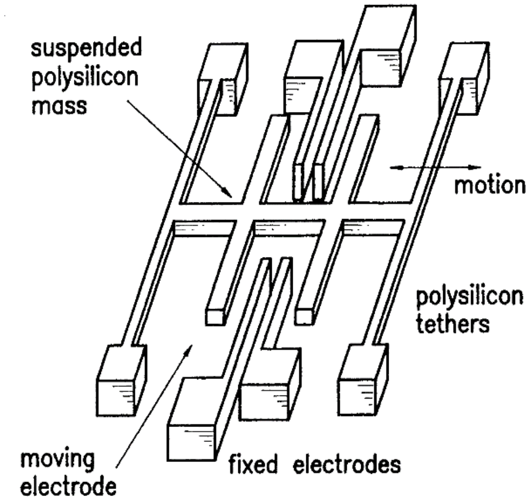
Kalman Gain: $K_{t+1|t} \coloneqq C_{t+1|t} S_{t+1|t}^{-1}$

$$m_{t+1|t} = \sum_{i=0}^{2n} W_i^{(m)} h\left(\mathcal{X}_{t+1|t}^{(i)}\right) \qquad \mathcal{X}_{t+1|t}^{(0)} = \mu_{t+1|t} \qquad \mathcal{X}_{t+1|t}^{(i)} = \mu_{t+1|t} \pm \sqrt{n+\lambda}\left[\sqrt{\Sigma_{t+1|t}}\right]_i$$

$$S_{t+1|t} = \sum_{i=0}^{2n} W_i^{(c)} \left(h\left(\mathcal{X}_{t+1|t}^{(i)}\right) - m_{t+1|t}\right)\left(h\left(\mathcal{X}_{t+1|t}^{(i)}\right) - m_{t+1|t}\right)^T + V$$

$$C_{t+1|t} = \sum_{i=0}^{2n} W_i^{(c)} \left(\mathcal{X}_{t+1|t}^{(i)} - \mu_{t+1|t}\right)\left(h\left(\mathcal{X}_{t+1|t}^{(i)}\right) - m_{t+1|t}\right)^T$$

MEMS Strapdown IMU



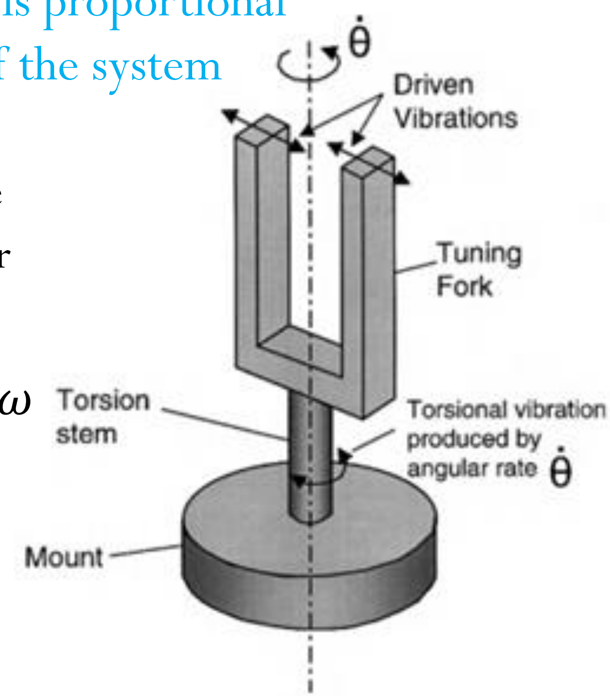
Surface Micromachined Accelerometer

- **MEMS:** micro-electro-mechanical system
- **IMU**: inertial measurement unit – consist of a triaxial accelerometer (measures linear acceleration) and a triaxial gyroscope (measures angular velocity)
- **Strapdown**: the IMU and the object/vehicle inertial frames are joined together/identical
- **Accelerometer**: a mass m on a spring with spring constant k

$$F = ma = kd \quad \Rightarrow \quad d = \frac{ma}{k}$$

Spring displacement is proportional to the acceleration of the system

  - VLSI Fabrication: the displacement of a metal plate with mass m is measured with respect to another plate using capacitance
  - Used for car airbags (if the acceleration goes above $2g$, the car is hitting something!)



- **Gyroscope**: uses Coriolis force to detect rotational velocity ω $(\dot{R} = R\widehat{\omega})$. The rotational velocity changes the mechanical resonance of a tuning fork
- **Homework**: Read Kraft's paper on quaternion-based UKF (**Canvas**)